AIAA JOURNAL
Vol. 33, No. 10, October 1995

Engineering Application of Experimental Uncertainty Analysis

Hugh W. Coleman*
University of Alabama in Huntsville, Huntsville, Alabama 35899
and
W. Glenn Steele†
Mississippi State University, Mississippi State, Mississippi 39762

Publication in late 1993 by the International Organization for Standardization (ISO) of the *Guide to the Expression of Uncertainty in Measurement* in the name of ISO and six other international organizations has, in everything but name only, established a new international experimental uncertainty standard. In this article, an analysis of the assumptions and approximations used in the development of the methods in the ISO guide is presented, and a comparison of the resulting equation with previously published uncertainty analysis approaches is made. Also discussed are the additional assumptions necessary to achieve the less complex "large sample" methodology that is recommended in AIAA Standard S-071-1995, *Assessment of Wind Tunnel Data Uncertainty*, issued in 1995. It is shown that these assumptions are actually less restrictive than those associated with some previously accepted methodologies. The article concludes with a discussion of some practical aspects of implementing experimental uncertainty analysis in engineering testing.

I. Introduction

DURING the past decade, and particularly in the past two years, there have been major developments in the field of experimental uncertainty analysis (or "measurement uncertainty," as it is often called). Publication in late 1993 by the International Organization for Standardization (ISO) of the *Guide to the Expression of Uncertainty in Measurement*[1] in the name of ISO and six other international organizations has, in everything but name only, established a new international experimental uncertainty standard. The methodology (but not the complete terminology) of the ISO guide is being adopted by groups working on formulating or revising guidelines, procedures, and standards in the experimental uncertainty area—a NATO AGARD Working Group on Quality Assessment for Wind Tunnel Testing,[2] of which the first author was a member, adopted the methodology, as has the committee (of which the second author is a member), currently revising the American National Standards Institute (ANSI)/American Society of Mechanical Engineers (ASME) Standard on Measurement Uncertainty.[3]

The authors' objective in this article is to present an analysis of the assumptions and approximations used in the development of the methods in the ISO guide and a discussion of the assumptions necessary to achieve a less complex "large sample" methodology that is applicable to the vast majority of engineering testing. As will be seen, these assumptions are actually less restrictive than those associated with some previously accepted methodologies.[3-5] The large sample methodology is that recommended in the AGARD report,[2] the draft of the revision to Ref. 3, and the newly issued AIAA standard.[6]

In Sec. II of this article a brief overview of uncertainty analysis is presented; in Sec. III a derivation of the uncertainty propagation equation is presented with particular discussion of the assumptions and approximations made; in Sec. IV a comparison of the resulting equation with previously published uncertainty analysis approaches is presented; in Sec. V additional assumptions are discussed that lead to a less complex uncertainty analysis approach that is suitable for most engineering applications; and the article concludes in Sec. VI with a summary and a discussion of some practical aspects of implementing experimental uncertainty analysis in engineering testing.

Received June 3, 1994; revision received March 13, 1995; accepted for publication March 15, 1995.

*Eminent Scholar in Propulsion and Professor, Propulsion Research Center, Department of Mechanical and Aerospace Engineering. Senior Member AIAA.

†Professor and Head, Department of Mechanical Engineering. Senior Member AIAA.

II. Uncertainty Analysis Overview

The word accuracy is generally used to indicate the relative closeness of agreement between an experimentally determined value of a quantity and its true value. Error (δ) is the difference between the experimentally determined value and the truth, and thus as error decreases, accuracy is said to increase. Only in rare instances is the true value of a quantity known. Thus, one is forced to estimate error, and that estimate is called an uncertainty U. Uncertainty estimates are made at some confidence level—a 95% confidence estimate, for example, means that the true value of the quantity is expected to be within the $\pm U$ interval about the experimentally determined value 95 times out of 100.

As shown in Fig. 1a, total error δ can be considered to be composed of two components: a precision (random) component ϵ and a bias (systematic) component β. An error is classified as precision if it contributes to the scatter of the data; otherwise, it is a bias error. It is assumed that corrections have been made for all systematic errors whose values are known. The remaining bias errors are thus equally as likely to be positive as negative.

Suppose that we are making a number of measurements of the value of a variable X that is steady. The k and $k+1$ measurements are shown in Fig. 1a. Since the bias is a fixed error, it is the same for each measurement. However, the precision error will have a different value for each measurement. It then follows that the total error in each measurement will be different, since the total error is the sum of the bias error and precision error in a measurement.

If we continued to take measurements as previously described until we had a sample of N readings, more than likely as N approached infinity the data would behave as shown in Fig. 1b. The bias error would be given by the difference between the mean (average) value μ of the N readings and the true value of X, whereas the precision errors would cause the frequency of occurrence of the readings to be distributed about the mean value.

As an estimator of β, a bias limit B is defined.[7] A 95% confidence estimate is interpreted as the experimenter being 95% confident that the true value of the bias error, if known, would fall within $\pm B$. A useful approach to estimating the magnitude of a bias error is to assume that the bias error for a given case is a single realization drawn from some statistical parent distribution of possible bias errors. For example, suppose a thermistor manufacturer specifies that 95% of samples of a given model are within ± 1.0 C of a reference resistance-temperature (R-T) calibration curve supplied with the

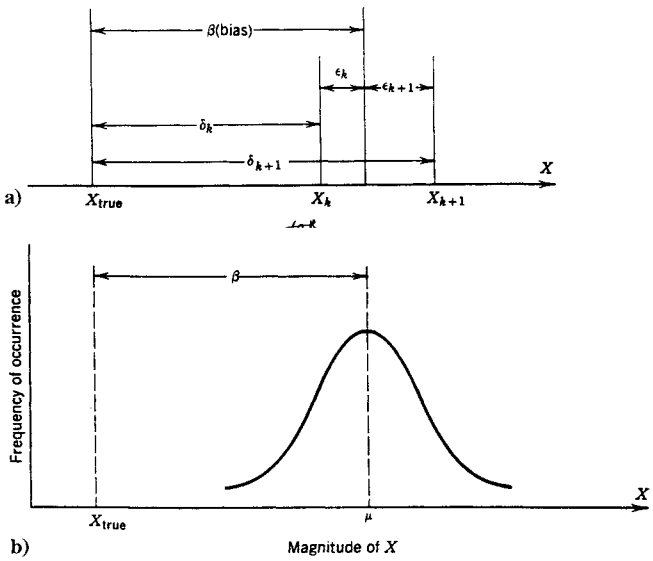

Fig. 1 Errors in the measurement of a variable X: a) two readings and b) infinite number of readings.

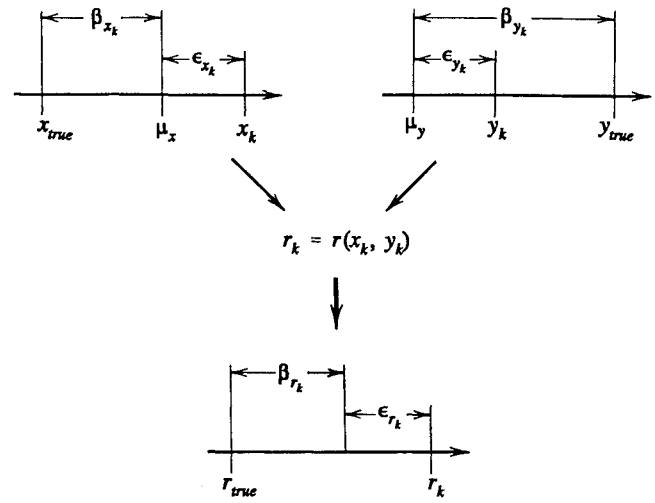

Fig. 2 Propagation of bias errors and precision errors into an experimental result.

thermistors. One might assume that the bias errors (the differences between the actual, but unknown, R-T curves of the various thermistors and the reference curve) belong to a Gaussian parent distribution with a standard deviation $b = 0.5$ C. Then the interval defined by $\pm B = \pm 2b = \pm 1.0$ C would include about 95% of the possible bias errors that could be realized from the parent distribution. (The bias limit is sometimes called the "systematic uncertainty.")

As an estimator of the magnitude of the precision errors (the width of the distribution of readings in Fig. 1b) a precision limit P is defined.[7] A 95% confidence estimate of P is interpreted to mean that the $\pm P$ interval about a single reading of X_i should cover μ 95 times out of 100. (The precision limit is sometimes called the "precision uncertainty.")

In nearly all experiments, the measured values of different variables are combined using a data reduction equation (DRE) to form some desired result. A good example is the experimental determination of drag coefficient of a particular model configuration in a wind-tunnel test. Defining drag coefficient as

$$C_D = \frac{2F_D}{\rho V^2 A} \tag{1}$$

one can envision that errors in the values of the variables on the right-hand side of Eq. (1) will cause errors in the experimental result C_D.

A more general representation of a data reduction equation is

$$r = r(X_1, X_2, \ldots, X_J) \tag{2}$$

where r is the experimental result determined from J measured variables X_i. Each of the measured variables contains bias errors and precision errors. These errors in the measured values then propagate through the data reduction equation, thereby generating the bias and precision errors in the experimental result r. In the following section, a derivation of the equation describing the error propagation is presented, with particular attention paid to the assumptions and approximations that are made to obtain the final equation.

III. Derivation of the Uncertainty Propagation Equation

Rather than present the derivation for the case in which the result is a function of many variables, the simpler case in which the result is a function of only two variables will be considered first. The expressions for the more general case will then be presented as extensions of the two-variable case.

Suppose that the data reduction equation is

$$r = r(x, y) \tag{3}$$

where the function is continuous and has continuous derivatives in the domain of interest. The situation is shown schematically in Fig. 2 for the kth set of measurements (x_k, y_k) that is used to determine r_k. Here, β_{x_k} and ϵ_{x_k} are the bias and precision errors, respectively, in the kth measurement of x, with a similar convention for the errors in y and in r. Assume that the test instrumentation and/or apparatus is changed for each measurement so that different values of β_{x_k} and β_{y_k} will occur for each measurement. Therefore, the bias errors and precision errors will be random variables so that

$$x_k = x_{\text{true}} + \beta_{x_k} + \epsilon_{x_k} \tag{4}$$

$$y_k = y_{\text{true}} + \beta_{y_k} + \epsilon_{y_k} \tag{5}$$

Now approximate the function r in the DRE using a Taylor series expansion. Expanding to the general point r_k from r_{true} gives

$$r_k = r_{\text{true}} + \frac{\partial r}{\partial x}(x_k - x_{\text{true}}) + \frac{\partial r}{\partial y}(y_k - y_{\text{true}}) + R_2 \tag{6}$$

where R_2 is the remainder term and where the partial derivatives are evaluated at $(x_{\text{true}}, y_{\text{true}})$. Since the true values of x and y are unknown, an approximation is always introduced when the derivatives are evaluated at some measured values (x_k, y_k).

The remainder term has the form[8]

$$R_2 = \frac{1}{2!}\left[\frac{\partial^2 r}{\partial x^2}(x_k - x_{\text{true}})^2 + 2\frac{\partial^2 r}{\partial x \partial y}(x_k - x_{\text{true}}) \times (y_k - y_{\text{true}}) + \frac{\partial^2 r}{\partial y^2}(y_k - y_{\text{true}})^2\right] \tag{7}$$

where the partial derivatives are evaluated at (ζ, χ) which is somewhere between (x_k, y_k) and $(x_{\text{true}}, y_{\text{true}})$. This term is usually assumed to be negligible, and so it is useful to consider the conditions under which this assumption might be reasonable. The factors $(x_k - x_{\text{true}})$ and $(y_k - y_{\text{true}})$ are the total errors in x and y. If the derivatives are of reasonable magnitude and the total errors in x and y are small, then R_2, containing the squares of the errors, will approach zero more quickly than will the first-order terms. Also, if $r(x, y)$ is a linear function, then the partial derivatives in Eq. (7) are identically zero (as is R_2).

Neglecting R_2, the expansion gives [taking r_{true} to the left-hand side (LHS)]

$$(r_k - r_{\text{true}}) = \frac{\partial r}{\partial x}(x_k - x_{\text{true}}) + \frac{\partial r}{\partial y}(y_k - y_{\text{true}}) \tag{8}$$

This expression relates the total error δ in the kth determination of the result r to the total errors in the measured variables and using the notation

$$\theta_x = \frac{\partial r}{\partial x} \tag{9}$$

can be written as

$$\delta_{r_k} = \theta_x\left(\beta_{x_k} + \epsilon_{x_k}\right) + \theta_y\left(\beta_{y_k} + \epsilon_{y_k}\right) \tag{10}$$

We are interested in obtaining a measure of the distribution of the various δ_r for (some large number) N determinations of the result r. The variance of the parent distribution is defined by

$$\sigma_{\delta_r}^2 = \lim(N \to \infty)\left[\frac{1}{N}\sum_{k=1}^{N}\left(\delta_{r_k}\right)^2\right] \tag{11}$$

Substituting Eq. (10) into Eq. (11) but deferring taking the limit gives

$$\begin{aligned}
&\frac{1}{N}\sum_{k=1}^{N}\left(\delta_{r_k}\right)^2 \\
&= \theta_x^2\frac{1}{N}\sum_{k=1}^{N}\left(\beta_{x_k}\right)^2 + \theta_y^2\frac{1}{N}\sum_{k=1}^{N}\left(\beta_{y_k}\right)^2 + 2\theta_x\theta_y\frac{1}{N}\sum_{k=1}^{N}\beta_{x_k}\beta_{y_k} \\
&+ \theta_x^2\frac{1}{N}\sum_{k=1}^{N}\left(\epsilon_{x_k}\right)^2 + \theta_y^2\frac{1}{N}\sum_{k=1}^{N}\left(\epsilon_{y_k}\right)^2 + 2\theta_x\theta_y\frac{1}{N}\sum_{k=1}^{N}\epsilon_{x_k}\epsilon_{y_k} \\
&+ 2\theta_x^2\frac{1}{N}\sum_{k=1}^{N}\beta_{x_k}\epsilon_{x_k} + 2\theta_y^2\frac{1}{N}\sum_{k=1}^{N}\beta_{y_k}\epsilon_{y_k} \\
&+ 2\theta_x\theta_y\frac{1}{N}\sum_{k=1}^{N}\beta_{x_k}\epsilon_{y_k} + 2\theta_x\theta_y\frac{1}{N}\sum_{k=1}^{N}\beta_{y_k}\epsilon_{x_k}
\end{aligned} \tag{12}$$

Taking the limit as N approaches infinity and using definitions of variances similar to that in Eq. (11), we obtain

$$\sigma_{\delta_r}^2 = \theta_x^2\sigma_{\beta_x}^2 + \theta_y^2\sigma_{\beta_y}^2 + 2\theta_x\theta_y\sigma_{\beta_x\beta_y} + \theta_x^2\sigma_{\epsilon_x}^2 + \theta_y^2\sigma_{\epsilon_y}^2 + 2\theta_x\theta_y\sigma_{\epsilon_x\epsilon_y} \tag{13}$$

assuming that there are no bias error/precision error correlations so that in Eq. (12) the final four terms containing $\beta\epsilon$ products are zero.

Since in reality we never know the various σ exactly, we must use estimates of them. Defining u_c^2 as an estimate of the variance of the distribution of total errors in the result, b^2 as an estimate of the variance of a bias error distribution, and S^2 as an estimate of the variance of a precision error distribution, we can write

$$u_c^2 = \theta_x^2 b_x^2 + \theta_y^2 b_y^2 + 2\theta_x\theta_y b_{xy} + \theta_x^2 S_x^2 + \theta_y^2 S_y^2 + 2\theta_x\theta_y S_{xy} \tag{14}$$

In Eq. (14) b_{xy} is an estimate of the covariance of the bias errors in x and the bias errors in y that is defined exactly by

$$\sigma_{\beta_x\beta_y} = \lim(N \to \infty)\left[\frac{1}{N}\sum_{k=1}^{N}\beta_{x_k}\beta_{y_k}\right] \tag{15}$$

Likewise, S_{xy} is an estimate of the covariance of the precision errors in x and y. In keeping with the nomenclature of the ISO guide,[1] u_c is called the combined standard uncertainty. For the more general case in which the experimental result is determined from Eq. (2), u_c is given by

$$\begin{aligned}
u_c^2 &= \sum_{i=1}^{J}\theta_i^2 b_i^2 + 2\sum_{i=1}^{J-1}\sum_{k=i+1}^{J}\theta_i\theta_k b_{ik} + \sum_{i=1}^{J}\theta_i^2 S_i^2 \\
&+ 2\sum_{i=1}^{J-1}\sum_{k=i+1}^{J}\theta_i\theta_k S_{ik}
\end{aligned} \tag{16}$$

where $(b_i)^2$ is the estimate of the variance of the bias error distribution of variable X_i, etc. The derivation to this point has been presented by the authors in Ref. 9 and in Appendix B of Ref. 7.

No assumptions about type(s) of error distributions are made to obtain the preceding equation for u_c. To obtain an uncertainty U_r (termed the expanded uncertainty in the ISO guide) at some specified confidence level (95%, 99%, etc.), the combined standard uncertainty u_c must be multiplied by a coverage factor K,

$$U_r = Ku_c \tag{17}$$

It is in choosing K that assumptions about the type(s) of the error distributions must be made.

An argument is presented in the ISO guide[1] that the error distribution of the result r may often be considered Gaussian because of the Central Limit Theorem, even if the error distributions of the X_i are not normal. In fact, the same argument can be made for approximate normality of the error distributions of the X_i since the errors typically are composed of a combination of errors from a number of elemental sources.

If it is assumed that the error distribution of the result r is normal, then the value of K for $C\%$ coverage corresponds to the $C\%$ confidence level t value from the t distribution (Table 1) so that

$$\begin{aligned}
U_r^2 &= \sum_{i=1}^{J}\theta_i^2(tb_i)^2 + 2\sum_{i=1}^{J-1}\sum_{k=i+1}^{J}\theta_i\theta_k\left(t^2 b_{ik}\right) + \sum_{i=1}^{J}\theta_i^2(tS_i)^2 \\
&+ 2\sum_{i=1}^{J-1}\sum_{k=i+1}^{J}\theta_i\theta_k\left(t^2 S_{ik}\right)
\end{aligned} \tag{18}$$

The effective number of degrees of freedom ν_r for determining the t value is given (approximately) by the so-called Welch–Satterthwaite formula as[1]

$$\nu_r = \frac{\left(\sum_{i=1}^{J}\left[\theta_i^2 S_i^2 + \theta_i^2 b_i^2\right]\right)^2}{\sum_{i=1}^{J}\left\{\left[(\theta_i S_i)^4/\nu_{si}\right] + \left[(\theta_i b_i)^4/\nu_{bi}\right]\right\}} \tag{19}$$

where the ν_{si} are the number of degrees of freedom associated with the S_i and the ν_{bi} are the number of degrees of freedom to associate with the b_i.

If an S_i has been determined from N_i readings of X_i taken over an appropriate interval, the number of degrees of freedom is given by

$$\nu_{si} = N_i - 1 \tag{20}$$

Table 1 Two-tailed t distribution giving t values for a confidence level C and number of degrees of freedom ν

ν	$C = 95.0\%$	$C = 99.0\%$
1	12.706	63.657
2	4.303	9.925
3	3.182	5.841
4	2.776	4.604
5	2.571	4.032
6	2.447	3.707
7	2.365	3.499
8	2.306	3.355
9	2.262	3.250
10	2.228	3.169
11	2.201	3.106
12	2.179	3.055
13	2.160	3.012
14	2.145	2.977
15	2.131	2.947
16	2.120	2.921
17	2.110	2.898
18	2.101	2.878
19	2.093	2.861
20	2.086	2.845
25	2.060	2.787
30	2.042	2.750
60	2.000	2.660
∞	1.960	2.576

For the number of degrees of freedom ν_{bi} to associate with a nonstatistical estimate of b_i, it is suggested in the ISO guide[1] that one might use the approximation

$$\nu_{bi} \approx \frac{1}{2}\left(\frac{\Delta b_i}{b_i}\right)^{-2} \tag{21}$$

where the quantity in parentheses is the relative uncertainty of b_i. For example, if one thought that the estimate of b_i was reliable to within ±25%, then

$$\nu_{bi} \approx \tfrac{1}{2}(0.25)^{-2} \approx 8 \tag{22}$$

Consideration of the 95% confidence t values in Table 1 reveals that the value of t approaches 2.0 (approximately) as the number of degrees of freedom increases. We thus face the somewhat paradoxical situation that as we have more information (ν_r increases), we can take t equal to 2.0 and do not have to deal with Eq. (19), but for the cases in which we have little information (ν_r small) we need to make the more detailed estimates required by Eq. (19). In Sec. V of this article, we examine the assumptions required to discard Eq. (19) and to simply use t equal to 2 (for 95% confidence).

IV. Comparison with Published Approaches

Approach of Abernethy et al.[5]

An approach that was widely used in the 1970s and 1980s was the U_{RSS}, U_{ADD} technique formulated by Abernethy et al.[5] and used in Refs. 3 and 4 and other Society of Automotive Engineers (SAE), Instrument Society of America (ISA), JANNAF, National Research Council (NRC), U.S. Air Force (USAF), NATO, and ISO standards documents.[5] According to Abernethy et al.,

$$U_{RSS} = \left[B_r^2 + (tS_r)^2\right]^{\frac{1}{2}} \tag{23}$$

for a 95% confidence estimate and

$$U_{ADD} = B_r + (tS_r) \tag{24}$$

for a 99% confidence estimate, where B_r is given by

$$B_r = \left[\sum_{i=1}^{J}\left(\theta_i^2 B_i^2\right)\right]^{\frac{1}{2}} \tag{25}$$

and B_r and the various B_i are 95% confidence bias limit estimates, and

$$S_r = \left[\sum_{i=1}^{J}\left(\theta_i^2 S_i^2\right)\right]^{\frac{1}{2}} \tag{26}$$

where t is the 95% confidence t value from the t distribution for ν_r degrees of freedom given by

$$\nu_r = \frac{(\theta_i S_i)^4}{\sum_{i=1}^{J}\left[(\theta_i S_i)^4/\nu_{si}\right]} \tag{27}$$

Consideration of these expressions in the context of the derivation presented in the previous section shows that they cannot be justified on a rigorous basis. The U_{ADD} approach has always been advanced on the basis of ad hoc arguments and with results from a few Monte Carlo simulations, but (as argued in the ISO guide[1]) for a 99% confidence level, the t value appropriate for 99% confidence should be used as the value of K in Eq. (17) to obtain a 99% confidence estimate for an assumed Gaussian distribution. The Abernethy et al. approaches also ignore the possibility of correlated bias error effects [taken into account in the b_{ik} covariance terms in Eq. (18)], although Ref. 3 does consider this effect in one example.

Approach of Coleman and Steele[7,9]

The authors, expanding on the ideas advanced by Kline and McClintock[10] and assuming Gaussian error distributions, proposed viewing Eq. (14) as a propagation equation for 68% confidence intervals. A 95% coverage estimate of the uncertainty in the result was then proposed as that given by an equation similar to Eq. (23)

$$U_r^2 = B_r^2 + P_r^2 \tag{28}$$

with the bias limit of the result defined by

$$B_r^2 = \sum_{i=1}^{J}\theta_i^2 B_i^2 + 2\sum_{i=1}^{J-1}\sum_{k=i+1}^{J}\theta_i\theta_k\rho_{bik}B_i B_k \tag{29}$$

and the precision limit of the result given by

$$P_r^2 = \sum_{i=1}^{J}\theta_i^2 (P_i)^2 + 2\sum_{i=1}^{J-1}\sum_{k=i+1}^{J}\theta_i\theta_k\rho_{sik}(P_i)(P_k) \tag{30}$$

where ρ_{bik} is the correlation coefficient appropriate for the bias errors in X_i and X_k, and ρ_{sik} is the correlation coefficient appropriate for the precision errors in X_i and X_k. The precision limit of the variable X_i is given by

$$P_i = t_i S_i \tag{31}$$

where t_i is determined with $\nu_i = N_i - 1$ degrees of freedom.

Equations (29) and (30) were viewed as propagation equations for 95% confidence bias limits and precision limits, and thus this approach avoided use of the Welch–Satterthwaite formula. Comparison of uncertainty coverages for a range of sample sizes using this approach and the Abernethy et al. approach have been previously presented.[11] As was stated earlier in reference to the Abernethy et al. approach, consideration of these expressions in the context of the derivation presented in the previous section shows that they cannot be justified on a rigorous basis.

Both the approach of Coleman and Steele[7] and the Abernethy et al. U_{RSS} approach[5] (properly modified to account for correlated bias effects) agree with the 95% confidence form of Eq. (18) for "large sample sizes"—that is, N_i (and ν_r) large enough so that t can be taken as 2.0. As will be discussed in Sec. V, an argument can be made that this corresponds roughly to $N_i \geq 10$.

Approach of ISO Guide[1]

The ISO guide[1] was published in late 1993 in the name of seven international organizations: the Bureau International des Poids et Mesures (BIPM), the International Electrotechnical Commission (IEC), the International Federation of Clinical Chemistry (IFCC), the International Organization for Standardization (ISO), the International Union of Pure and Applied Chemistry (IUPAC), the International Union of Pure and Applied Physics (IUPAP), and the International Organization of Legal Metrology (OIML). It is now, in the authors' opinion, the de facto international standard.

One fundamental difference between the approach of the guide and that of Eqs. (16), (18), and (19) in this article is that the guide uses $u(x_i)$, a "standard uncertainty," to represent the quantities b_i and S_i used in this article. Instead of categorizing uncertainties as either bias (systematic) or precision (random), the various u are divided into type A standard uncertainties and type B standard uncertainties. Type A uncertainties are those evaluated "by the statistical analysis of series of observations," whereas type B uncertainties are those evaluated "by means other than the statistical analysis of series of observations." These do not correspond to the categories described by the traditional engineering usage: precision (or random or repeatability) uncertainty and bias (or systematic or fixed) uncertainty.

Arguments can, of course, be made for both sets of nomenclature. Types A and B unambiguously define how an uncertainty estimate was made, whereas systematic and precision uncertainties can change from one category to the other in a given experimental program depending on the experimental process used—a systematic calibration uncertainty can become a source of scatter (and thus a precision uncertainty) if a new calibration is done before each reading in the sample is taken, for example. On the other hand, if one

wants an estimate of the expected dispersion of results for a particular experimental approach or process, then the systematic/precision categorization is useful—particularly when used in the "debugging" phase of an experiment.[7]

Considering the tradition in engineering of the bias/precision (systematic/random) uncertainty categorization and its usefulness in engineering experimentation as mentioned earlier, this article's authors have chosen to retain that categorization while adopting the mathematical procedures of the ISO guide. This categorization is also used in the AGARD[2] document, by the committee currently revising Ref. 3, and in the AIAA standard.[6]

Approach of NIST[12]

Taylor and Kuyatt[12] reported guidelines for the implementation of a National Institute of Standards and Technology (NIST) policy that states the following:

> Use expanded uncertainty U to report the results of all NIST measurements other than those for which u_c has traditionally been employed. To be consistent with current international practice, the value of k to be used at NIST for calculating U is, by convention, $k = 2$. Values of k other than 2 are only to be used for specific applications dictated by established and documented requirements.

(The coverage factor k corresponds to the K used in this article.) The NIST approach is thus that in the ISO guide,[1] and no confidence level is associated with U when reported by NIST even though the coverage factor is specified as 2.0.

Approaches of AGARD[2] and AIAA Standard[6]

The first author was a participant in AGARD Fluid Dynamics Panel Working Group 15 on Quality Assessment for Wind Tunnel Testing and was the principal author of the methodology chapter in the resulting report.[2] This AGARD report, with minor revisions, has been issued in 1995 as an AIAA standard.[6] The recommended methodology is that discussed in this article, including the additional assumptions discussed later in Sec. V so that t is taken as 2 "unless there are other overriding considerations."[2]

V. Additional Assumptions for Engineering Applications

In much engineering testing (such as in most wind-tunnel tests, for example) it seems that the use of the preceding complex but still approximate equations [Eqs. (18) and (19)] for U_r and ν_r derived in Sec. III of this article would be excessively and unnecessarily complicated and would tend to give a false sense of the degree of significance of the numbers computed using them.[2,6] The following discussion examines what additional simplifying approximations can reasonably be made for application of uncertainty analysis in most engineering testing.

Approximating the Coverage Factor

Consider the process of estimating the uncertainty components (tb_i) and (tS_i) and obtaining U_r. The propagation equation and the Welch–Satterthwaite formula are approximate, not exact, and the Welch–Satterthwaite formula does not include the influence of correlated uncertainties. In addition, unavoidable uncertainties are always present in estimating the bias uncertainties b_i and in estimating their associated degrees of freedom, ν_{bi}.

In fact, the uncertainty associated with an S_i calculated from N readings of X_i can be surprisingly large.[1] For samples from a Gaussian parent population with standard deviation σ, 95 out of 100 determinations of S_i will scatter within an interval of approximately $\pm 0.45\sigma$ if the S_i are determined from $N = 10$ readings and within an interval of approximately $\pm 0.25\sigma$ if the S_i are determined from $N = 30$ readings. (A sample with $N > 30$ has traditionally been considered a "large" sample.[3]) This effect seems to have received little consideration in the engineering "measurement uncertainty" literature. As stated in the ISO guide,[1]

> This . . . shows that the standard deviation of a statistically estimated standard deviation is not negligible for practical values of n. One may therefore conclude that Type A evaluations of standard uncertainty are not necessarily more reliable than Type B evaluations, and that in many practical measurement situations where the number of observations is limited, the components obtained from Type B evaluations may be better known than the components obtained from Type A evaluations.

Considering the 95% confidence t table (Table 1), one can observe that for $\nu_r \geq 9$ the values of t are within about 13% of the large sample t value of 2. This difference is relatively insignificant compared with the uncertainties inherent in estimating the S_i and b_i as discussed earlier. Therefore, for most engineering applications it is proposed that Gaussian error distributions and $\nu_r \geq 9$ be assumed so that $t = 2$ always. (This could be called the "large sample assumption.") This assumption eliminates the need for evaluation of ν_r using the Welch–Satterthwaite formula and thus the need to estimate all of the ν_{xi} and the ν_{bi}.

Consideration of the Welch–Satterthwaite formula [Eq. (19)] shows that, because of the exponent of 4 in each term, ν_r is most influenced by the number of degrees of freedom of the largest of the $\theta_i S_i$ or $\theta_i b_i$ terms. If, for example, $\theta_3 S_3$ is dominant, then $\nu_r \approx \nu_{s3} \geq 9$ for $N_3 \geq 10$ [recalling Eq. (20)]. If, on the other hand, $\theta_3 b_3$ is dominant, then $\nu_r \approx \nu_{b3} \geq 9$ when the relative uncertainty in b_i is about 24% or less [recalling Eq. (21)]. Therefore, invoking the large sample assumption essentially means that if a $\theta_i S_i$ is dominant, then its $N_i \geq 10$, or if a $\theta_i b_i$ is dominant, then the relative uncertainty in that b_i is about 24% or less. If there is no single dominant term, but there are M different $\theta_i S_i$ and $\theta_i b_i$ that all have the same magnitude and same number of degrees of freedom ν_a, then

$$\nu_r = M\nu_a \tag{32}$$

If $M = 3$, for example, ν_a would only have to be 3 or greater for ν_r to be equal to or greater than 9. Therefore, t can often legitimately be taken as 2 for estimating the uncertainty in a result determined from several measured variables even when the numbers of degrees of freedom associated with the measured variables are very small.

If the large sample assumption is made so that $t = 2$, then the 95% confidence expression for U_r becomes

$$U_r^2 = \sum_{i=1}^{J} \theta_i^2 (2b_i)^2 + 2\sum_{i=1}^{J-1}\sum_{k=i+1}^{J} \theta_i\theta_k (2)^2 (b_{ik}) + \sum_{i=1}^{J} \theta_i^2 (2S_i)^2 + 2\sum_{i=1}^{J-1}\sum_{k=i+1}^{J} \theta_i\theta_k (2)^2 (S_{ik}) \tag{33}$$

Recalling the definitions of bias limit and precision limit from Sec. II, the $(2b_i)$ factors are equal to the 95% confidence bias limits B_i. The $2S_i$ factors, however, do not necessarily correspond to the 95% confidence precision limits, as discussed later. If no further assumptions are made, we can rewrite Eq. (33) as

$$U_r^2 = \sum_{i=1}^{J} \theta_i^2 B_i^2 + 2\sum_{i=1}^{J-1}\sum_{k=i+1}^{J} \theta_i\theta_k B_{ik} + \sum_{i=1}^{J} \theta_i^2 (2S_i)^2 + 2\sum_{i=1}^{J-1}\sum_{k=i+1}^{J} \theta_i\theta_k (2)^2 (S_{ik}) \tag{34}$$

where B_{ik} is the 95% confidence estimator of the covariance of the bias errors in X_i and X_k.

Estimating Precision Limits

Using a Monte Carlo type simulation, the behavior of the precision limit determined from N readings of a variable X was investigated as a function of N. A Gaussian parent population with mean μ and standard deviation σ was specified for a variable X. For $N = 8$, for example, the results were calculated as follows. Eight readings

were randomly drawn from the specified parent population, and a sample standard deviation was calculated using

$$S = \left[\frac{1}{N-1} \sum_{k=1}^{N} [X_k - \bar{X}]^2 \right]^{\frac{1}{2}} \tag{35}$$

where the mean value is defined as

$$\bar{X} = \frac{1}{N} \left[\sum_{k=1}^{N} X_k \right] \tag{36}$$

A check was then made to see if the $\pm(tS)$ and $\pm(2S)$ intervals centered about the next (ninth) randomly chosen reading covered μ, and the appropriate counter was incremented when the check was positive. This procedure was followed for 100,000 trials, and the percent coverage was the percentage of the 100,000 trials in which μ was covered. An average, S_{avg}, of the 100,000 S was calculated, and the precision limit ratios $(2S_{avg}/2\sigma)$ and $(tS_{avg}/2\sigma)$ were determined. This entire procedure was followed for N from 3 to 30, and the results are shown in Figs. 3 and 4.

For $N = 10$, the percent coverage by $2S$ is 92.5% and the average $2S$ interval is about 0.98 of the 2σ interval, whereas the percent coverage by tS is 95% and the average tS interval is about 1.10 of the 2σ interval. As N increases, the percent coverage by $2S$ approaches 95% and the average $2S$ interval approaches the 2σ interval value. Considering the previous discussion of the dispersion associated with determinations of S using N readings, it is proposed that the 95% confidence precision limit for a variable X_i be estimated as

$$\begin{aligned} P_i &= t_i S_i \qquad N < 10 \\ P_i &= 2S_i \qquad N \geq 10 \end{aligned} \tag{37}$$

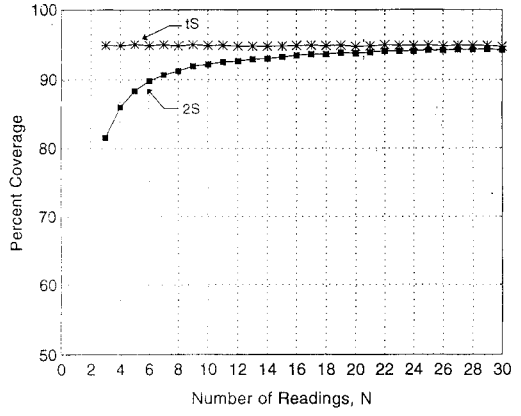

Fig. 3 Percent coverage of population mean μ for N readings of a Gaussian-distributed variable.

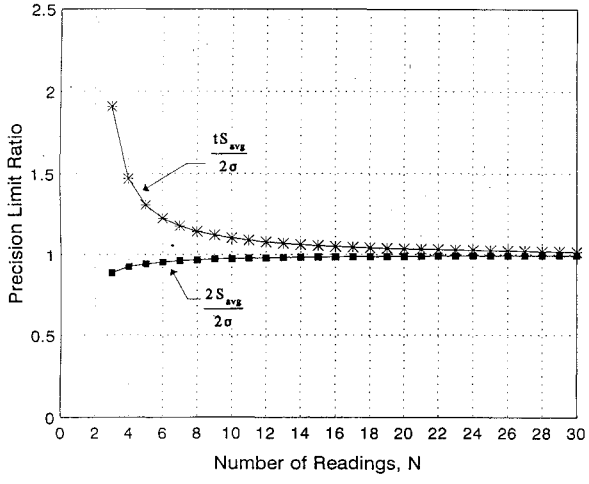

Fig. 4 Ratio of calculated to true precision limit for N readings of a Gaussian-distributed variable.

where t_i is determined with $N_i - 1$ degrees of freedom.

In cases in which all $N_i \geq 10$, then the various $(2S_i)$ in Eq. (34) correspond to the precision limits P_i so that

$$U_r^2 = \sum_{i=1}^{J} \theta_i^2 B_i^2 + 2 \sum_{i=1}^{J-1} \sum_{k=i+1}^{J} \theta_i \theta_k B_{ik} + \sum_{i=1}^{J} \theta_i^2 P_i^2 + 2 \sum_{i=1}^{J-1} \sum_{k=i+1}^{J} \theta_i \theta_k P_{ik} \tag{38}$$

where P_{ik} equals $4(S_{ik})$.

If we define the bias limit (systematic uncertainty) of the result as

$$B_r^2 = \sum_{i=1}^{J} \theta_i^2 B_i^2 + 2 \sum_{i=1}^{J-1} \sum_{k=i+1}^{J} \theta_i \theta_k B_{ik} \tag{39}$$

and the precision limit (precision uncertainty) of the result as

$$P_r^2 = \sum_{i=1}^{J} \theta_i^2 P_i^2 + 2 \sum_{i=1}^{J-1} \sum_{k=i+1}^{J} \theta_i \theta_k P_{ik} \tag{40}$$

then Eq. (38) can be written as

$$U_r^2 = B_r^2 + P_r^2 \tag{41}$$

and Eqs. (39) and (40) can be viewed as propagation equations for the bias limits and precision limits, respectively.

Single Test

When the result is determined from a single test, that is, at a given test condition the result is determined once using Eq. (2)

$$r = r(X_1, X_2, \ldots, X_J) \tag{2}$$

and when the various X_i are considered single measurements, then Eq. (40) is used to find the precision limit of the result. This situation is often encountered in large-scale engineering tests in which measurements of the variables are made at a given set point over a period that is small compared with the periods of the factors causing variability in the experiment. A proper precision limit (one indicative of the dispersion of the variable over several cycles of the factors causing its variation) cannot be calculated from readings taken over such a small time interval. For such data, the measurement(s) of a variable X_i should be considered a single reading—whether the value of X_i is the average of 10, 10^3, or 10^6 readings taken during the short measurement time. In such a test, the value for the precision limit to be associated with a single reading would have to be based on previous information about that measurement obtained over the appropriate time interval.[11] If previous readings of a variable over an appropriate interval are not available, then the experimenter must estimate a value for P_i using the best information available at that time.[2,7]

For single tests in which some of the variables (X_2 and X_3, for instance) can be determined as averages from multiple readings over an appropriate time period but the other variables cannot be, then

$$r = r(X_1, \bar{X}_2, \bar{X}_3, \ldots, X_J) \tag{42}$$

and Eq. (40) is used to find the precision limit of the result as follows. For the variables that are single readings, the various P_i are the precision limits determined from previous information or estimated from the best available information. For the averaged variables when N_2 and N_3 are equal to or greater than 10, P_2 and P_3 should be taken as precision limits of means, $(2S_2)/(N_2)^{1/2}$ and $(2S_3)/(N_3)^{1/2}$, with the various S calculated using Eq. (35). When N_2 and N_3 are less than 10, it is the authors' recommendation that the precision limits used in Eq. (40) for the averaged variables be taken as $(P_2)/(N_2)^{1/2}$ and $(P_3)/(N_3)^{1/2}$, where P_2 and P_3 are determined from previous information, as is done for the single reading variables.

For tests in which multiple readings of all of the variables can be obtained over an appropriate period, the following method is recommended.

Multiple Tests

If a test is performed so that M multiple sets of measurements $(X_1, X_2, \ldots, X_J)_k$ at the same test condition are obtained, then M results can be determined using Eq. (2) and an average result $\bar{r}$ can be determined using

$$\bar{r} = \frac{1}{M}\sum_{k=1}^{M} r_k \tag{43}$$

If the M sets of measurements were obtained over an appropriate time period, the precision limit that should be associated with a single result would be

$$P_r = tS_r \tag{44}$$

where t is determined with $M - 1$ degrees of freedom [and as in Eq. (37) is taken as 2 for $M \geq 10$] and S_r is the standard deviation of the sample of M results

$$S_r = \left[\frac{1}{M-1}\sum_{k=1}^{M}(r_k - \bar{r})^2\right]^{\frac{1}{2}} \tag{45}$$

The precision limit that should be associated with the average result is given by

$$P_{\bar{r}} = \frac{P_r}{\sqrt{M}} \tag{46}$$

with P_r given by Eq. (44). The uncertainty that should be associated with $\bar{r}$ is (using the large sample assumption)

$$U_{\bar{r}}^2 = B_r^2 + \left(2S_r/\sqrt{M}\right)^2 \tag{47}$$

with B_r given by Eq. (39).

Correlated Precision Uncertainties

The S_{ik} terms in Eq. (34) and the P_{ik} terms in Eq. (38) take into account the possibility of precision errors in different variables being correlated. These terms have traditionally been neglected,[3–5,7,10] although precision errors in different variables caused by the same uncontrolled factor(s) are certainly possible and can have a substantial impact on the value of the precision limit.[13] In such cases, one would need to acquire sufficient data to allow a valid statistical estimate of the precision covariance terms to be made if using Eqs. (34) or (38). Note, however, that the multiple tests approach using Eq. (47) implicitly includes the correlated error effect—a definite advantage when multiple sets of measurements over an appropriate time period are available.

Estimating Bias Limits

Bias Limits of Individual Variables

When attempting to estimate the bias limits B_i of the individual variables in Eqs. (34) or (39), one might separate the bias errors that influence the measurement of a variable into different categories: calibration errors, data acquisition errors, data reduction errors, test technique errors, etc. Within each category, there may be several elemental sources of bias. For instance, if for the Jth variable, X_J, there are M elemental bias errors identified as significant and whose bias limits are estimated as $(B_J)_1, (B_J)_2, \ldots, (B_J)_M$, then the bias limit for the measurement of X_J is calculated as the root-sum-square (rss) combination of the elemental limits

$$B_J = \left[\sum_{k=1}^{M}(B_J)_k^2\right]^{\frac{1}{2}} \tag{48}$$

The elemental bias limits, $(B_i)_k$, must be estimated for each variable X_i using the best information one has available at the time. In the design phase of an experimental program, manufacturer's specifications, analytical estimates, and previous experience will typically provide the basis for most of the estimates. As the experimental program progresses, equipment is assembled, and calibrations are conducted, these estimates can be updated using the additional information gained about the accuracy of the calibration standards, errors associated with the calibration process and curvefit procedures, and perhaps analytical estimates of installation errors.

As Moffat[14] suggests, there can be additional conceptual bias errors resulting from not measuring the variable whose symbol appears in the data reduction equation. An example would be a freestream velocity value measured at a particular axial position on a wind-tunnel centerline and used as the freestream velocity at that cross section in determining C_D, but there may be a cross-sectional gradient of velocity at that location causing the average value to be different. Hence, the inclusion of an elemental bias term for the conceptual error would be appropriate.

Examples of estimating elemental bias limits in engineering experiments are presented in Refs. 2, 6, 7, and 15, and the procedures for estimating bias uncertainties assumed from a non-Gaussian distribution are presented in Ref. 1.

Correlated Bias Limits

Correlated bias limits are those that are not independent of each other, typically a result of different measured variables sharing some identical elemental error sources. It is not unusual for the uncertainties in the results of experimental programs to be influenced by the effects of correlated bias errors in the measurements of several of the variables. A typical example occurs when different variables are measured using the same transducer, such as multiple pressures sequentially ported to and measured with the same transducer or temperatures at different positions in a flow measured with a single probe that is traversed across the flowfield. Obviously, the bias errors in the variables measured with the same transducer are not independent of one another. Another common example occurs when different variables are measured using different transducers all of which have been calibrated against the same standard, a situation typical of the electronically scanned pressure (ESP) measurement systems in wide use in aerospace test facilities. In such a case, at least a part of the bias error arising from the calibration procedure will be the same for each transducer, and thus some of the elemental bias error contributions in the measurements of the variables will be correlated.

The B_{ik} terms in Eqs. (34) or (39) must be approximated—there is in general no way to obtain the data with which to make a statistical estimate of the covariance of the bias errors in X_i and the bias errors in X_j. The approximation of such terms was considered in detail in Ref. 16, where it was shown that the approach that consistently gives the most satisfactory approximation for the correlated bias limits was

$$B_{ik} = \sum_{\alpha=1}^{L}(B_i)_\alpha (B_k)_\alpha \tag{49}$$

where L is the number of elemental systematic error sources that are common for measurements of variables X_i and X_k.

If, for example,

$$r = r(X_1, X_2) \tag{50}$$

and it is possible for portions of the bias limits B_1 and B_2 to arise from the same source(s), then Eq. (39) gives

$$B_r^2 = \theta_1^2 B_1^2 + \theta_2^2 B_2^2 + 2\theta_1\theta_2 B_{12} \tag{51}$$

For a case in which the measurements of X_1 and X_2 are each influenced by four elemental error sources and sources 2 and 3 are the same for both X_1 and X_2, Eq. (48) gives

$$B_1^2 = (B_1)_1^2 + (B_1)_2^2 + (B_1)_3^2 + (B_1)_4^2 \tag{52}$$

and

$$B_2^2 = (B_2)_1^2 + (B_2)_2^2 + (B_2)_3^2 + (B_2)_4^2 \tag{53}$$

whereas Eq. (49) gives

$$B_{12} = (B_1)_2(B_2)_2 + (B_1)_3(B_2)_3 \tag{54}$$

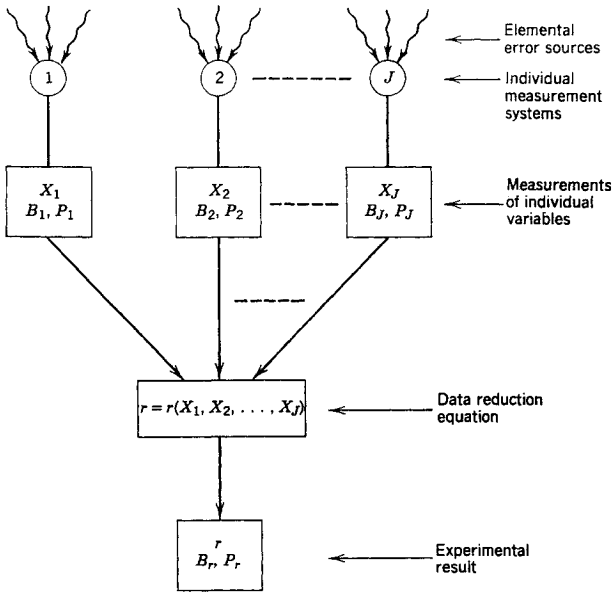

Fig. 5 Propagation of uncertainties into an experimental result.

VI. Summary

The 95% confidence large sample uncertainty analysis approach recommended for use in the vast majority of engineering testing can be summarized as follows.

In single tests with precision limits estimated using previous information and all $N_i \geq 10$,

$$U_r^2 = B_r^2 + P_r^2 \tag{55}$$

The bias limit of the result is

$$B_r^2 = \sum_{i=1}^{J} \theta_i^2 B_i^2 + 2 \sum_{i=1}^{J-1} \sum_{k=i+1}^{J} \theta_i \theta_k B_{ik} \tag{56}$$

where B_{ik} is given by Eq. (49). The precision limit of the result is given by

$$P_r^2 = \sum_{i=1}^{J} \theta_i^2 P_i^2 \tag{57}$$

assuming no correlated precision uncertainties.

In multiple tests, for which there are M results at the same experimental set point, the uncertainty that should be associated with $\bar{r}$ is

$$U_{\bar{r}}^2 = B_r^2 + \left(2S_r/\sqrt{M}\right)^2 \tag{58}$$

with B_r given by Eq. (56) and S_r by Eq. (45).

Experimental uncertainty analysis using the large sample approach can be implemented in an engineering test as described next. (The authors emphasize that the partial derivatives in the equations do not have to be determined analytically; they can be numerically approximated using a finite difference scheme, for example, that perturbs the data reduction equation and computes a numerical value for each θ_i.)

1) For each experimental result, the data reduction equation [Eq. (2)] should be identified; once this has been done, the various X_i that must be considered are known. How well each result needs to be known (what U_r is acceptable?) should be decided in the initial planning phase of the experimental program, and this should be used to guide decisions made in the phases that follow.

2) Those variables that are the most important from an uncertainty viewpoint should be identified. The authors suggest[7] initially assuming all uncertainties are random and using Eq. (57) in a general uncertainty analysis to investigate the sensitivity of the uncertainty in the result to the uncertainties in the variables. This analysis is sometimes done in a parametric study using an assumed range of uncertainties over the range of anticipated test conditions. Once this study has been accomplished, resources can be concentrated, if necessary, on obtaining detailed estimates of the uncertainties in the most important variables.

3) A detailed uncertainty analysis[7] should then be performed considering both systematic uncertainties and precision uncertainties. In doing this, it is sometimes useful to visualize the experiment and the propagation of uncertainties as shown schematically in Fig. 5. For each X_i, estimates of the bias limit and the precision limit are usually made.

It is generally easiest to obtain an estimate of the bias limit for X_i by estimating the bias limits of the significant elemental sources and combining them using Eq. (48). To achieve this, the sources of systematic uncertainty for each X_i should be identified by considering the potential elemental error sources affecting the determination of that variable. Once the sources of uncertainty have been identified, their relative significance should be established. This is often done using order of magnitude estimates of the sources. As a rule of thumb for a given X_i, those uncertainty sources that are smaller than one-fourth or one-fifth of the largest sources can usually be considered negligible. Resources can then be concentrated on obtaining good estimates of those uncertainties of most importance.

In the authors' experience, for most engineering testing it is generally not cost effective or necessary to try to estimate precision limits at the elemental error source level. (The authors have observed cases in which significant resources were squandered through attempts to obtain good, large sample estimates of elemental precision uncertainties that were of absolutely no consequence in the test program.) It is generally far more effective to estimate the precision limits of the measured variables (at the P_i level in Fig. 5) and obtain P_r using Eq. (57) or, if possible, computing the direct estimate of S_r as given by Eq. (45) if multiple results at the same set point are available (perhaps from previous tests of a similar type).

4) The overall uncertainty, bias limit, and precision limit for the experimental result r are then found using the preceding equations.

5) Each experimental result that is reported should be accompanied by its overall uncertainty, reference to an explanation of the estimation of its precision and bias limits, and a discussion of any implications of the uncertainty on the use that should be made of the result.

Acknowledgments

The authors gratefully acknowledge the discussions with and comments from their faculty colleagues and their students, the members of NATO/AGARD Working Group 15, the members of the ASME Performance Test Codes committee revising Ref. 3, and the many people who reviewed drafts of Refs. 2 and 6. Kendall K. Brown performed the Monte Carlo analysis that produced the results shown in Figs. 3 and 4.

References

[1] Anon., *Guide to the Expression of Uncertainty in Measurement*, International Organization for Standardization, Geneva, Switzerland, 1993, ISBN 92-67-10188-9.

[2] Anon., *Quality Assessment for Wind Tunnel Testing*, AGARD-AR-304, July 1994.

[3] Anon., *Measurement Uncertainty*, American National Standards Institute/American Society of Mechanical Engineers, PTC 19.1-1985 Pt. 1, American Society of Mechanical Engineers, New York, 1986.

[4] Anon., *Measurement Uncertainty for Fluid Flow in Closed Conduits*, American National Standards Institute/American Society of Mechanical Engineers, MFC-2M-1983, American Society of Mechanical Engineers, New York, 1984.

[5] Abernethy, R. B., Benedict, R. P., and Dowdell, R. B., "ASME Measurement Uncertainty," *Journal of Fluids Engineering*, Vol. 107, June 1985, pp. 161–164.

[6] Anon., *Assessment of Wind Tunnel Data Uncertainty*, AIAA Standard S-071-1995, Washington, DC, June 1995.

[7] Coleman, H. W., and Steele, W. G., *Experimentation and Uncertainty Analysis for Engineers*, Wiley, New York, 1989.

[8] Hildebrand, F. B., *Advanced Calculus for Applications*, Prentice–Hall, Englewood Cliffs, NJ, 1962.

[9] Coleman, H. W., and Steele, W. G., "Some Considerations in the Propagation of Bias and Precision Errors into an Experimental Result," *Experimental Uncertainty in Fluid Measurements*, FED Vol. 58, American Society of Mechanical Engineers, New York, 1987, pp. 57–62.

[10]Kline, S. J., and McClintock, F. A., "Describing Uncertainties in Single-Sample Experiments," *Mechanical Engineering*, Vol. 75, Jan. 1953, pp. 3–8.

[11]Steele, W. G., Taylor, R. P., Burrell, R. E., and Coleman, H. W., "The Use of Data from Previous Experience to Estimate the Precision Uncertainty of Small Sample Experiments," *AIAA Journal*, Vol. 31, No. 10, 1993, pp. 1891–1896.

[12]Taylor, B. N., and Kuyatt, C. E., "Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results," National Inst. of Standards and Technology, NIST TN 1297, Gaithersburg, MD, Sept. 1994.

[13]Hudson, S. T., Bordelon, W., and Coleman, H. W., "Effect of Correlated Precision Errors on the Uncertainty of a Subsonic Venturi Calibration," AIAA Paper 95-0797, Jan. 1995.

[14]Moffat, R. J., "Describing the Uncertainties in Experimental Results," *Experimental Thermal and Fluid Science*, Vol. 1, No. 1, 1988, pp. 3–17.

[15]Coleman, H. W., Hosni, M. H., Taylor, R. P., and Brown, G. B., "Using Uncertainty Analysis in the Debugging and Qualification of a Turbulent Heat Transfer Test Facility," *Experimental Thermal and Fluid Science*, Vol. 4, No. 6, 1991, pp. 673–683.

[16]Brown, K. K., Coleman, H. W., Steele, W. G., and Taylor, R. P., "Evaluation of Correlated Bias Approximations in Experimental Uncertainty Analysis," AIAA Paper 94-0772, Jan. 1994.